Knowledge Distillation in Iterative Generative Models for Improved Sampling Speed

Year: 2021 | Citations: 339 | Authors: Eric Luhman, Troy Luhman

---

Abstract

Iterative generative models, such as noise conditional score networks and denoising diffusion probabilistic models, produce high quality samples by gradually denoising an initial noise vector. However, their denoising process has many steps, making them 2-3 orders of magnitude slower than other generative models such as GANs and VAEs. In this paper, we establish a novel connection between knowledge distillation and image generation with a technique that distills a multi-step denoising process into a single step, resulting in a sampling speed similar to other single-step generative models. Our Denoising Student generates high quality samples comparable to GANs on the CIFAR-10 and CelebA datasets, without adversarial training. We demonstrate that our method scales to higher resolutions through experiments on 256 x 256 LSUN. Code and checkpoints are available at https://github.com/tcl9876/Denoising_Student